

RECEIVED

FEB 19 2003

TECH CENTER 1600/2900



1600

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/612,809B

DATE: 02/10/2003

TIME: 15:04:09

Input Set : A:\Iowa042.app

Output Set : N:\CRF4\02102003\I612809B.raw

3 <110> APPLICANT: SHEFFIELD, VAL C.  
 4 AIWARI, WALLACE L.M.  
 5 STONE, EDWIN M.  
 6 HISHIMURA, DARRYL  
 7 PATIL, SHIVA  
 9 <120> TITLE OF INVENTION: THERAPEUTICS AND DIAGNOSTICS FOR CONGENITAL HEART  
 10 DISEASE BASED ON A NOVEL HUMAN TRANSCRIPTION FACTOR  
 12 <130> FILE REFERENCE: IOWA:042USDI  
 14 <140> CURRENT APPLICATION NUMBER: 09/612,809B  
 15 <141> CURRENT FILING DATE: 2000-07-10  
 17 <160> NUMBER OF SEQ ID NOS: 20  
 19 <170> SOFTWARE: PatentIn Ver. 2.1  
 21 <210> SEQ ID NO: 1  
 22 <211> LENGTH: 2284  
 23 <212> TYPE: DNA  
 24 <213> ORGANISM: Homo sapiens  
 26 <240> SEQUENCE: 1

ENTERED

```

17 cgaataaaag taccggggggg cccgggggagg cggccggggc gggggccccc cccccccccc 60
18 cccgtttatg taccggggggg cccgggggagg cccgggggagg ggtctctgga agggcgaggc 120
19 aggtgggaga gaagggggggg tgggtgttct tcttttttgt ctggtttccc cgttttgccc 180
20 cgggaagatg cccgggggagg tctgggaagg cgggtctggc cggccggggg cgggtctctc 240
21 cctctgggag gaccggggggg cggggggggg cggggggggg ggttagccga gggcgaggag 300
22 gaggaagggg caggagaggg cgggagaggg gggagaggg cgggaaggg agggcaggg 360
23 ggggggagga ggtggggggg gggggggggg gggggggggg gggggggggg cggggggggg 420
24 cccggggggg ggtggggggg gggggggggg gggggggggg gggggggggg gggggggggg 480
25 cccgaagggg aatcaaatcg aaccccaagg caggaaaagg taaaaggacc catcaaggga 540
26 aatcggaagg caaaaaaaa aatcccaatt aaaaaaaa cctgagaata tccaccacac 600
27 cagggaatcg aatatacctc caaaaattca gctcaccagg accagcaaga agaaaaactc 660
28 attttcttaa cagattaatt cagagccacc tccactttgc cttgtctaaa taaccacacc 720
29 cgtaaaactg tctatacaga gacagcaaaa ccttggttta ttaaaggaca ggtttactcc 780
30 agataacagc taagttttct cttggttttc agagaccggc tttccctccc tcccggtccc 840
31 cctctcttgc cctctctctc gctctccacc tgttaagatat tattttctcc tatgttgaag 900
32 ggagggggaa agtcccccgt tatgaaagtc gttttctctc taccgatgga cttgttttaa 960
33 aatgtaaatg gcaacatagt aatttatttt taatttgtag ttggatgtcg tggacacacc 1020
34 gccagaaaag gttcccaaaa cctgacgtta aatttgctga aattttaaat tctggttttt 1080
35 ttctcattat aaaaagggaa actgatttaa tcttattcta tctcttttcc tttctttttg 1140
36 ttgataatg tttgtttg tttattata aatttacctt cagtgtgaat gagactata 1200
37 cgtctggata ctttaataga gttttaatta ttacgaaaaa agatttcaga gataaaaaac 1260
38 cagaagttac ctattctcca ctttaatttc tgaataatgg agaaaccttc tgaactagtc 1320
39 atgtcaaatc ctactcaaaa ctttttttgt tagatttatt tccctgcagg atctttctga 1380
40 aatgttacta tatagtcagg cttggtttgag gctagtaaaa agatattttt ctaaaacagat 1440
41 tggagtgggg atataaaca atactgtttc tcaataatga cagtcacatg tccggaaaatt 1500
42 ttaagcccat gaatcagcgg cgggtctacc accgtgatgc ctgtgtgccc agagatggga 1560

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/612,809B

DATE: 02/10/2003

TIME: 15:04:09

Input Set : A:\Iowa042.app

Output Set: N:\CRF4\02102003\I612809B.raw

```

53 ctgtgogggac agatatgcac agataaatat ttggcttggtg tattccatat aaaattgcag 1620
54 tgratattat acatccctgt gagccagatg ctgaatagat ttttccctat tattccagtc 1680
55 ctttataaaa ggaaaaataa accagttttt aaatgtatgt atataattct ccccattht 1740
56 caatccctca tgtattacat agaaggattg cttttttaaa aatatactgc ggggttgaaa 1800
57 gggatattta atctttgaga aactatttta gaaaatatgt ttgtagaaca attatttttg 1860
58 aaaaagattt aaagcaataa caagaaggaa ggcgagagga gcagaacatt ttggtctagg 1920
59 gtggtttctt tttaaaccaat tttttcttgt taatttacag ttaaacctag gggacaatcc 1980
60 ggahtggccc tccctctttt gtaataaacc caggaaatgt aataaattca ttatcttagg 2040
61 tttttctgac ctgccaatca gactttgggg agatggcgat ttgattacag acgttcgggg 2100
62 gggggggggg cttgcagttt gttttggaga taatacagtt tctgtctatc tggcgctct 2160
63 atctagaggg aacacttaag cagtaattgc tgttgcttgt tgtcaaat ttgatattgt 2220
64 aaaggattg ctgcaataa atacacttta atttcagtc aaaaaaaaaa aaaaaaaaaa 2280
65 aaaa
66 C10 - SEQ ID NO: 2
67 C11 - LENGTH: 553
68 C12 - TYPE: PRO
69 C13 - ORGANISM: Homo sapiens
70 C100 - SEQUENCE: 2
71 Met Gln Ala Arg Tyr Ser Val Ser Ser Pro Asn Ser Leu Gly Val Val
72 1 5 10 15
73 Pro Tyr Leu Gly Glu Gln Ser Tyr Tyr Arg Ala Ala Ala Ala Ala
74 20 25 30
75 Ala Gly Gly Gly Tyr Thr Ala Met Pro Ala Pro Met Ser Val Tyr Ser
76 35 40 45
77 His Pro Ala His Ala Glu Gln Tyr Pro Gly Gly Met Ala Arg Ala Tyr
78 50 55 60
79 Gly Pro Tyr Thr Pro Gln Pro Gln Pro Lys Asp Met Val Lys Pro Pro
80 65 70 75 80
81 Tyr Ser Tyr Ile Ala Leu Ile Thr Met Ala Ile Gln Asn Ala Pro Asp
82 85 90 95
83 Lys Lys Ile Thr Leu Asn Gly Ile Tyr Gln Phe Ile Met Asp Arg Phe
84 100 105 110
85 Pro Phe Tyr Arg Asp Asn Lys Gln Gly Trp Gln Asn Ser Ile Arg His
86 115 120 125
87 Asn Leu Ser Leu Asn Glu Cys Phe Val Lys Val Pro Arg Asp Asp Lys
88 130 135 140
89 Lys Pro Gly Lys Gly Ser Tyr Trp Thr Leu Asp Pro Asp Ser Tyr Asn
90 145 150 155 160
91 Met Phe Glu Asn Gly Ser Phe Leu Arg Arg Arg Arg Arg Phe Lys Lys
92 165 170 175
93 Lys Asp Ala Val Lys Asp Lys Glu Glu Lys Asp Arg Leu His Leu Lys
94 180 185 190
95 Glu Pro Pro Pro Pro Gly Arg Gln Pro Pro Pro Ala Pro Pro Glu Gln
96 195 200 205
97 Ala Asp Gly Asn Ala Pro Gly Pro Gln Pro Pro Pro Val Arg Ile Gln
98 210 215 220
99 Asp Ile Lys Thr Glu Asn Gly Thr Cys Pro Ser Pro Pro Gln Pro Leu
100 225 230 235 240
101 Ser Pro Ala Ala Ala Leu Gly Ser Gly Ser Ala Ala Val Pro Lys

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/612,809B

DATE: 02/10/2003

TIME: 15:04:09

Input Set : A:\Iowa042.app

Output Set: N:\CRF4\02102003\I612809B.raw

```

120                               245                               250                               155
122 Ile Glu Ser Pro Asp Ser Ser Ser Ser Ser Leu Ser Ser Gly Ser Ser
123                               260                               270
125 Pro Pro Gly Ser Leu Pro Ser Ala Arg Pro Leu Ser Leu Asp Gly Ala
126                               275                               280                               285
128 Asp Ser Ala Pro Pro Pro Pro Ala Pro Ser Ala Pro Pro Pro His His
129                               290                               295                               300
131 Ser Gln Gly Phe Ser Val Asp Asn Ile Met Thr Ser Leu Arg Gly Ser
132                               305                               310                               315
134 Pro Gln Ser Ala Ala Ala Glu Leu Ser Ser Gly Leu Leu Ala Ser Ala
135                               320                               325                               330
137 Ala Ala Ser Ser Arg Ala Gly Ile Ala Pro Pro Leu Ala Leu Gly Ala
138                               335                               340                               345
140 Tyr Ser Pro Gly Gln Ser Ser Leu Tyr Ser Ser Pro Cys Ser Gln Thr
141                               350                               355                               360
143 Ser Ser Ala Gly Ser Ser Gly Gly Gly Gly Gly Gly Ala Gly Ala Ala
144                               365                               370                               375
146 Gly Gly Ala Gly Gly Ala Gly Thr Tyr His Cys Asn Leu Glu Ala Met
147                               380                               385                               390
149 Ser Leu Tyr Ala Ala Gly Glu Arg Gly Gly His Leu Gln Gly Ala Pro
150                               395                               400                               405
152 Gly Gly Ala Gly Gly Ser Ala Val Asp Asn Pro Leu Pro Asp Tyr Ser
153                               410                               415                               420
155 Leu Pro Pro Val Thr Ser Ser Ser Ser Ser Ser Ser Leu Ser His Gly Gly
156                               425                               430                               435
158 Gly Gly Gly Gly Gly Gly Gly Gly Gln Glu Ala Gly His His Pro Ala
159                               440                               445                               450
161 Ala His Gln Gly Arg Leu Thr Ser Trp Tyr Leu Asn Gln Ala Gly Gly
162                               455                               460                               465
164 Asp Leu Gly His Leu Ala Ser Ala Ala Ala Ala Ala Ala Ala Gly
165                               470                               475                               480
167 Tyr Pro Gly Gln Gln Gln Asn Phe His Ser Val Arg Glu Met Phe Glu
168                               485                               490                               495
170 Ser Gln Arg Ile Gly Leu Asn Asn Ser Pro Val Asn Gly Asn Ser Ser
171                               500                               505                               510
173 Cys Gln Met Ala Phe Pro Ser Ser Gln Ser Leu Tyr Arg Thr Ser Gly
174                               515                               520                               525
176 Ala Phe Val Tyr Asp Cys Ser Lys Phe
177                               530                               535                               540
179                               545                               550
180 (Q10) SEQ ID NO: 3
181 (Q11) LENGTH: 1662
182 (Q12) TYPE: DNA
183 (Q13) ORGANISM: Homo sapiens
184 (Q14) SEQUENCE: 3
185 atgcaggcgc gctactccgc gtcagccccc aaactccctgg gattggtgac ctactccggc 60
186 ggagagagaa gctactacgc cggggcgggc cgggcgggc ggaggggcta caacggccatg 120
187 cgggccccca ttagcgtgta ctgcacccct jcggaagcag agaggtaccc ggggggcatg 180
188 gcccgcgcct acggggccca cagcccgag ccggaaccca aggacatggt gaaagccccc 240
189 tatactaca tcgggtcat caccatggcc atccagaacg cccggacaa gaagatcacc 300

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/612,809B

DATE: 02/10/2003

TIME: 15:04:09

Input Set : A:\Iowa042.app

Output Set: N:\CRF4\02102003\I612809B.raw

```

191 ctgaacgggca tctaacagtt catcatagac cgtttccctt tctaccggga caa-aagagag 360
192 gctgggacaa acagatctcg ccacaaactc tggctcaacg agtgccttct caa-gggggcg 420
193 caggacacaa agaaacgggq caaggggagc taatggacgc tggaccggga ctcctacaaac 480
194 atgttcacaa agggagcttt cctggcgagg cggcgagcgt tcaagaagaa ggaacgggag 540
195 caggacaaag aggaagagaa cagctcgac ctcaagggag cgcacccgac cggcgggag 600
196 ccccgagcag cggcgaggaa gcaaggcgac gcaaacggc cgggtccgaa gggcggggac 660
197 atggccttcc aggaatctaa gaacgaagac ggtacgtgca cctggcggac cgaagcctcg 720
198 tcccggggag cggccttgag caggggagc ggcggcggcg tgcacaaagt cgaagagcgc 780
199 cccagcgcaa ggaagagcct gtccagggg agcagacccc cgggcagcct ggcgtgggag 840
200 cggcgctaa gctgggaggg tggggattcc ggcagcgc cgcggcgac ctccggcgg 900
201 cggcgcaaac atagtcaggg attcagcgag gacaaatcca tgacgtcgtt gggggggctg 960
202 cggcagaggg cggcggggaa gctcagctcc ggccttatgg cctggcgac cgggtcctcg 1020
203 cggcggggaa tggcaacccc gctggcgctc ggccctact cggccggcga gacgtccctc 1080
204 tggacctccc cctggagcga gacctcagc gggggagct cgggggggag cgggggggag 1140
205 cggcggggag cgggggggag ggggggggag gggacatacc actgcaactt gcaagcagat 1200
206 cactggaag cggcggggga ggcgggggag cacttgacag ggcggcccg gggcggggag 1260
207 cctcggggag tggcaaaccc cctggcgga cactctctgc ctccgggtac cagcagcagc 1320
208 cctcgctccc tggctcaggg cggcggggag gggcggggag gggcggggag gggcggggag 1380
209 cctcagctcg cggcggggag aggcggcctc acctcgtggt acctgaacca gggcggggag 1440
210 cactcgggg ccttgccaag cggcggggag gggcggggag cgcagagcta cccgggggag 1500
211 cagcagaaat tccactcgtt gggggagatg ttccagtcac agaggatcgg cttgaacaaac 1560
212 tctccagtaa cgggaatag tagctgtcaa atggcttcc cttcagcga gtctctgtac 1620
213 ccaacgtcg gaccttctgt ctacgactgt agcaagtttt ga 1682

```

216 &gt;10&gt; SEQ ID NO: 4

217 &gt;11&gt; LENGTH: 106

218 &gt;12&gt; TYPE: PRT

219 &gt;13&gt; ORGANISM: Artificial Sequence

220 &gt;100&gt; FEATURE:

221 &gt;130&gt; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

222 &gt;100&gt; Feature

223 &gt;400&gt; SEQUENCE: 4

```

224 1  Lys  Lys  Asp  Met  Val  Lys  Phe  Phe  Tyr  Ser  Tyr  Ile  Ala  Leu  Ile  Thr
225 5  10  15
226 20  Ala  Ile  Gln  Asn  Ala  Pro  Asp  Lys  Lys  Ile  Thr  Leu  Asn  Gly  Ile
227 30  35  40  45
228 50  Tyr  Gln  Phe  Ile  Met  Asp  Arg  Phe  Phe  Phe  Tyr  Arg  Asp  Asn  Lys  Gln
229 60  65  70  75  80
230 85  Gly  Trp  Gln  Asn  Ser  Ile  Arg  His  Asn  Leu  Ser  Leu  Asn  Glu  Cys  Phe
231 90  95  100
232 105  Lys  Val  Phe  Arg  Asp  Asp  Lys  Lys  Phe  Gly  Lys  Gly  Ser  Tyr  Trp
233 110  115  120  125  130
234 135  Thr  Leu  Asp  Phe  Asp  Ser  Tyr  Asn  Met  Phe  Glu  Asn  Gly  Ser  Phe  Leu
235 140  145  150  155  160  165
236 165  Arg  Arg  Arg  Arg  Arg  Phe  Lys  Lys  Lys  Asp
237 170  175  180  185

```

238 &gt;10&gt; SEQ ID NO: 5

239 &gt;11&gt; LENGTH: 106

240 &gt;12&gt; TYPE: PRT

241 &gt;13&gt; ORGANISM: Artificial Sequence

## RAW SEQUENCE LISTING

DATE: 02/10/2003

PATENT APPLICATION: US/09/612,809B

TIME: 15:04:09

Input Set : A:\Iowa042.app

Output Set: N:\CRF4\02102003\I612809B.raw

253 &lt;220&gt; FEATURE:

254 &lt;223&gt; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

255 Peptide

257 &lt;400&gt; SEQUENCE: 5

258 110 Lys Asp Leu Val Lys Pro Pro Tyr Ser Tyr Ile Ala Leu Ile Thr

259 1 5 10 15

260 201 Met Ala Ile Gln Asn Ala Pro Glu Lys Lys Ile Thr Leu Asn Gly Ile

261 20 15 30

262 304 Tyr Gln Phe Ile Met Asp Arg Phe Pro Phe Tyr Arg Glu Asn Lys Gln

263 35 40 45

264 367 Gly Trp Gln Asn Ser Ile Arg His Asn Leu Ser Leu Asn Glu Cys Phe

265 50 55 60

266 370 Val Lys Val Pro Arg Asp Asp Lys Lys Pro Gly Lys Gly Ser Tyr Trp

267 65 70 75 80

268 372 Thr Leu Asp Pro Asp Ser Tyr Asn Met Phe Glu Asn Gly Ser Phe Leu

269 85 90 95

270 374 Arg Arg Arg Arg Arg Phe Lys Lys Lys Asp

271 100 105

272 &lt;400&gt; SEQ ID NO: 6

273 &lt;410&gt; LENGTH: 106

274 &lt;412&gt; TYPE: PRT

275 &lt;413&gt; ORGANISM: Artificial Sequence

276 &lt;400&gt; FEATURE:

277 &lt;413&gt; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

278 Peptide

279 &lt;400&gt; SEQUENCE: 5

280 290 Thr Thr Glu Pro Thr Lys Pro Pro Tyr Ser Tyr Ile Ala Leu Ile Ala

281 1 5 10 15

282 301 Met Ala Ile Gln Ser Ser Pro Gly Gln Arg Ala Thr Leu Ser Gly Ile

283 20 15 30

284 304 Tyr Arg Val Ile Met Gly Arg Phe Ala Phe Tyr Arg His Asn Arg Pro

285 35 40 45

286 307 Gly Trp Gln Asn Ser Ile Arg His Asn Leu Ser Leu Asn Glu Cys Phe

287 50 55 60

288 311 Val Lys Val Pro Arg Asp Asp Arg Lys Pro Gly Lys Gly Ser Tyr Trp

289 65 70 75 80

290 312 Thr Leu Asp Pro Asp Cys His Asp Met Phe Glu His Gly Ser Phe Leu

291 85 90 95

292 314 Arg Arg Arg Arg Arg Phe Thr Arg Gln Thr

293 100 105

294 &lt;400&gt; SEQ ID NO: 7

295 &lt;410&gt; LENGTH: 106

296 &lt;412&gt; TYPE: PRT

297 &lt;413&gt; ORGANISM: Artificial Sequence

298 &lt;400&gt; FEATURE:

299 &lt;413&gt; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

300 Peptide

301 &lt;400&gt; SEQUENCE: 7

302 322 Ala Glu Thr Pro Gln Lys Pro Pro Tyr Ser Tyr Ile Ala Leu Ile Ala

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/612,809B

DATE: 02/10/2003

TIME: 15:04:10

Input Set : A:\Iowa042.app

Output Set: N:\CRF4\02102003\I612809B.raw